

AD-A077 096

WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER  
NEWTON'S METHOD FOR GENERALIZED EQUATIONS. (U)  
JUN 79 N H JOSEPHY  
MRC-TSR-1965

F/G 12/1

UNCLASSIFIED

DAA629-75-C-0024  
NL

1 OF 1  
AD-A077096



END  
DATE  
FILMED  
12-79  
DDC

AD A 077096

MRC Technical Summary Report #1965

NEWTON'S METHOD FOR GENERALIZED  
EQUATIONS

Norman H. Josephy

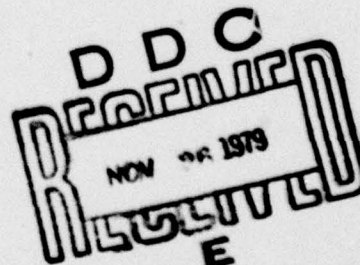
LEVEL

Mathematics Research Center  
University of Wisconsin-Madison  
610 Walnut Street  
Madison, Wisconsin 53706

June 1979

Received May 25, 1979

DDC FILE COPY



Approved for public release  
Distribution unlimited

Sponsored by

U. S. Army Research Office  
P.O. Box 12211  
Research Triangle Park  
North Carolina 27709

National Science Foundation  
Washington, D. C. 20550

UNIVERSITY OF WISCONSIN - MADISON  
MATHEMATICS RESEARCH CENTER

NEWTON'S METHOD FOR GENERALIZED EQUATIONS

Norman H. Josephy

Technical Summary Report #1965  
June 1979

ABSTRACT

Newton's method is a well known and often applied technique for computing a zero of a nonlinear function. By using the theory of generalized equations, a Newton method is developed to solve problems arising in both mathematical programming and mathematical economics.

~~We prove~~ <sup>are proved</sup> two results concerning the convergence and convergence rate of Newton's method <sub>A</sub> for generalized equations. Examples are given to emphasize the application of this method to generalized equations representing the non-linear programming problem and the nonlinear complementarity problem.

~~We present~~ <sup>are both presented</sup> computational results of Newton's method applied to a nonlinear complementarity problem of Kojima, [3] and an invariant capital stock problem of Hansen and Koopmans, [2].

AMS(MOS) Subject Classifications: 49D99, 90C99, 90A15

Key Words: Variational Inequalities, Complementarity, Newton's Method

Work Unit No. 5 - Mathematical Programming and Operations Research

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024. This material is based upon work supported by the National Science Foundation under Grant No. DCR 74-20584 and Grant No. MCS 74-20584 A02 and the Graduate School of the University of Wisconsin-Madison.



## Significance and Explanation

The solution to many practical problems in operations research and mathematical economics, such as the nonlinear programming problem and the economic equilibrium problem, can be found by solving a collection of equations and inequalities. Robinson [10] has shown that this collection of equations and inequalities can be written in a form very similar to an equation, which he calls a generalized equation. Extensive work by many researchers has created a well-developed theory and computational algorithms for solving equations. The analogy between an equation and a generalized equation can be utilized to extend this theory and algorithms to the more difficult problems involving equations and inequalities. In this paper, one of the most successful algorithms for computing the solution of an equation, Newton's method, is shown to be applicable to generalized equations. To illustrate how the method is applied to practical problems, two examples are solved using Newton's method. The first is a nonlinear complementarity problem used by Kojima [3] to test other algorithms. The second example is the invariant capital stock problem developed by Hansen and Koopmans [2]. Thus, the range of problems to which Newton's method can be applied has been extended to include finding solutions to nonlinear programming problems and equilibria in economic systems.

Accession For	
NTIS G.A.I.	<input checked="" type="checkbox"/>
DOC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	<input type="checkbox"/>
By	
Distribution/	
Availability Codes	
Avail and/or	
Dist	A

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.



## NEWTON'S METHOD FOR GENERALIZED EQUATIONS

Norman H. Josephy

1. Introduction. A generalized equation is a set-valued analogue of a single-valued equation. Many problems arising in mathematical programming and mathematical economics can be analyzed by generalizing, to the set-valued case, classical techniques used in the study of single-valued equations. Preliminary investigations of the properties and applications of generalized equations have been carried out by Robinson [6, 7, 8, 9, 10]. This paper is a contribution to this continuing research.

Newton's method is a classical technique of both theoretical and practical importance in the study and computation of solutions to single-valued equations. The theory supporting Newton's method develops a set of conditions under which the iterative procedure which comprises the method will converge, at a known rate, to a solution of an equation. Extensive computational experience supports the notion that under suitable conditions, this algorithm is an efficient computational procedure. A Newton method for generalized equations will extend to problems in mathematical programming and mathematical economics these desirable properties.

The definition and examples of generalized equations are given in section 2. The linearization of a generalized equation is defined in section 3, and is illustrated by applying it to the nonlinear complementarity problem and finding a Kuhn-Tucker point of a nonlinear programming problem. The condition of strong regularity is defined in section 4, and a corollary to a result of Robinson [9] is stated. The two main theorems of this paper, the domain of attraction theorem and the Newton-Kantorovich

---

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024. This material is based upon work supported by the National Science Foundation under Grant No. DCR 74-20584 and Grant No. MCS 74-20584 A02 and the Graduate School of the University of Wisconsin-Madison.

theorem, are stated and proved in section 5. We conclude with some computational results obtained by applying Newton's method to the generalized equations representing a nonlinear complementarity problem of Kojima [3] and an invariant capital stock problem of Hansen and Koopmans [2].



2. Generalized Equations. Let  $C$  be a non-empty, convex, closed subset of  $R^n$ . The normal cone to  $C$  at  $x \in C$  is the set of outward pointing normals to  $C$  at  $x$ . We have the following:

Definition 1. Let  $C$  be a non-empty, closed, convex subset of  $R^n$ .

The normal cone to  $C$  at  $x$  is given by

$$N_C(x) = \begin{cases} \{z \mid \langle z, k-x \rangle \leq 0 \ \forall k \in C\} & \text{if } x \in C, \\ \emptyset & \text{if } x \notin C. \end{cases}$$

We can now define a generalized equation.

Definition 2. Let  $f: D \subset R^n \rightarrow R^n$ .

Let  $C$  be a non-empty, closed, convex subset of  $R^n$ .

A generalized equation is a set relation

$$0 \in f(x) + N_C(x).$$

Definition 3. A solution  $x^*$  of the generalized equation

$$0 \in f(x) + N_C(x) \text{ is a vector } x^* \text{ satisfying}$$

$$0 \in f(x^*) + N_C(x^*).$$

Thus,  $x^*$  is a solution if and only if  $x^* \in C$  and  $\langle f(x^*), k-x^* \rangle \geq 0$  for all  $k \in C$ .

The following example shows the nonlinear complementarity problem as a special case of a generalized equation when the set  $C$  is chosen as a non-negative orthant.

Example 1. Let  $K$  be a non-empty, closed, convex cone in  $R^n$ .

Define the dual cone of  $K$ ,  $K^*$ , to be

$$K^* = \{y \mid \langle y, k \rangle \geq 0, \ \forall k \in K\}$$

Suppose  $x \in K$ ,  $y \in R^n$  and  $\langle y, k-x \rangle \geq 0$  for all  $k \in K$ . This implies, for  $k=2x$ , that  $\langle y, x \rangle \geq 0$ . But choosing  $k=\frac{1}{2}x$  yields  $\langle y, x \rangle \leq 0$ . Thus,  $\langle y, x \rangle = 0$ . Hence  $0 \leq \langle y, k \rangle - \langle y, x \rangle = \langle y, k \rangle$  for all  $k \in K$ , that is,  $y \in K^*$ . We have thus shown that if  $x \in K$  and  $\langle y, k-x \rangle \geq 0$  for all  $k \in K$ ,



then  $x \in K$ ,  $y \in K^*$ , and  $\langle x, y \rangle = 0$ . The converse holds trivially, that is, if  $x \in K$ ,  $y \in K^*$ , and  $\langle x, y \rangle = 0$ , then  $x \in K$  and  $\langle y, k-x \rangle \geq 0$ , for all  $k \in K$ . Using Definition 3, we have established the following equivalence:

$$0 \in y + N_K(x) \text{ if and only if } x \in K, y \in K^*, \text{ and } \langle x, y \rangle = 0.$$

For the special case of  $K = R_+^n$  and  $F: R^n \rightarrow R^n$ , we have that

$$0 \in F(x^*) + N_K(x^*) \text{ if and only if } x^* \geq 0, F(x^*) \geq 0, \text{ and } \langle x^*, F(x^*) \rangle = 0.$$

This is the nonlinear complementarity problem.

In the next example of this section, we derive the generalized equation which represents the Kuhn-Tucker conditions for a nonlinear programming problem. The result of Example 1 will be used with  $K$  defined as the Cartesian product of cones.

Example 2. Let  $f: R^n \rightarrow R$ ,  $g: R^n \rightarrow R^m$ , and  $h: R^n \rightarrow R^p$  be continuously differentiable maps. We will consider the following nonlinear programming problem.

$$(N.L.P.) : \min f(x) \text{ subject to } g(x) \leq 0, h(x) = 0.$$

We define the Lagrangian of (N.L.P.) as

$$L(x, u, w) = f(x) + \sum_{i=1}^m u_i g_i(x) + \sum_{j=1}^p w_j h_j(x) .$$

The gradient of  $L$  with respect to  $x$  will be denoted  $\nabla L$ . We say that  $(x^*, u^*, w^*)$  is a Kuhn-Tucker triple if and only if  $\nabla L(x^*, u^*, w^*) = 0$ ,  $u^* \geq 0$ ,  $g(x^*) \leq 0$ ,  $\langle u^*, g(x^*) \rangle = 0$ ,  $h(x^*) = 0$ . Let  $K := \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^p$ , the Cartesian product of cones. Note that the dual cone  $K^* = \{0\}^n \times \mathbb{R}_+^m \times \{0\}^p$ . Finally, define the function  $F$  on  $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$  to  $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$  by

$$F(x, u, w) = \begin{bmatrix} \nabla L(x, u, w) \\ -g(x) \\ -h(x) \end{bmatrix}$$

Then  $(x^*, u^*, w^*)$  is a Kuhn-Tucker triple if and only if  $(x^*, u^*, w^*)$  solves the generalized equation

$$0 \in F(x, u, w) + N_K(x, u, w) .$$

3. Linearization of Generalized Equations. The approximation of a nonlinear equation by a linear equation is a well known and frequently successful technique used in the study of equations. In this section, we define and illustrate the linear approximation of a generalized equation. We begin with a definition.

Definition 4. Let  $D$  be a non-empty, open, convex subset of  $\mathbb{R}^n$ .

Let  $F:D \rightarrow \mathbb{R}^n$  be a map with continuous derivative  $F'$ .

Fix  $\bar{x} \in D$ .

The linearization of  $F$  at  $\bar{x}$ , denoted by  $LF_{\bar{x}}$  is defined as

$$LF_{\bar{x}}(x) := F(\bar{x}) + F'(\bar{x})(x - \bar{x}) .$$

Thus, the linearization of  $F$  at  $\bar{x}$  is an affine map passing through the point  $\bar{x}$  having "slope" equal to that of  $F$  at  $\bar{x}$ . The linearization of a generalized equation at  $\bar{x}$  is obtained by replacing  $F$  with  $LF_{\bar{x}}$ .

Definition 5. Let  $F:D \rightarrow \mathbb{R}^n$  have continuous derivative  $F'$ .

The linearization of the generalized equation

$$0 \in F(x) + N_K(x)$$

at  $\bar{x}$  is the generalized equation



$$0 \in LF_{\bar{x}}(x) + N_K(x) .$$

We illustrate the linearization procedure with two examples. The linearization of the generalized equation for the nonlinear complementarity problem is given in Example 3. The result is a linear complementarity problem. This is analogous to the case of nonlinear equations, where linearization results in linear equations. The linearization of the generalized equation for the Kuhn-Tucker triple of a nonlinear programming problem is given in Example 4. The result is a generalized equation equivalent to the Kuhn-Tucker conditions for a quadratic programming problem. This quadratic program was first proposed by Wilson [11] as the subproblem in an iterative scheme to solve concave nonlinear programming problems. The first extensive analysis of the use of these linearizations to solve certain problems in both mathematical programming and mathematical economics can be found in Robinson [6]. In anticipation of the remarks in section 5, we note that the result of a linearization of a generalized equation will be the subproblem in a Newton method used to compute solutions to the given generalized equation.

Example 3. Consider the nonlinear complementarity problem in  $R^n$ :

$$F(x) \geq 0, \quad x \geq 0, \quad \langle x, F(x) \rangle = 0 .$$

By Example 1, this is equivalent to the generalized equation

$$0 \in F(x) + N_K(x) ,$$

where  $K: = \mathbb{R}_+^n$ . The linearization of this generalized equation at  $\bar{x}$  is

$$0 \in F(\bar{x}) + F'(\bar{x})(x - \bar{x}) + N_K(x) .$$

Regrouping the constant vectors and relabelling as follows:

$$a := F(\bar{x}) - F'(\bar{x})\bar{x} , M := F'(\bar{x})$$

the linearization can be rewritten as

$$0 \in a + Mx + N_K(x) ,$$

which is equivalent to the linear complementarity problem

$$a + Mx \geq 0 , x \geq 0 , \langle x, a + Mx \rangle = 0 .$$

Thus, linearization applied to the nonlinear complementarity problem results in a linear complementarity problem. A Newton's method for solving a nonlinear complementarity problem will then consist of iteratively solving a sequence of linear complementarity problems.

Example 4. We consider the linearization of the Kuhn-Tucker conditions for the nonlinear-programming problem, as given in Example 2. Before doing so, we will use the results of Example 2 to derive the generalized equation whose solution is a Kuhn-Tucker triple for the following quadratic programming problem.

$$(Q.P.) \quad \min \frac{1}{2}(x - \bar{x})^T H(x - \bar{x}) + (x - \bar{x})^T c \quad \text{subject to}$$

$$A(x - \bar{x}) + a \leq 0 \quad \text{and} \quad B(x - \bar{x}) + b = 0 ,$$



where  $\bar{x} \in R^n$  is fixed and  $H$  is a symmetrix  $n \times n$  matrix. Let  $D(x,u,w) := H(x-\bar{x}) + c + A^T u + B^T w$  be the gradient of the Lagrangian for (Q.P.). Using the results of Example 2, a Kuhn-Tucker triple must satisfy the generalized equation

$$0 \in F(x,u,w) + N_K(x,u,w),$$

where  $K := R^n \times R_+^m \times R^p$  and

$$(4) \quad F(x,u,w) := \begin{bmatrix} D(x,u,w) \\ -A(x-\bar{x})-a \\ -B(x-\bar{x})-b \end{bmatrix} = \begin{bmatrix} c \\ -a \\ -b \end{bmatrix} + \begin{bmatrix} H & A^T & B^T \\ -A & 0 & 0 \\ -B & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x-\bar{x} \\ u \\ w \end{bmatrix}$$

Now consider the nonlinear programming problem of Example 2.

(N.L.P.):  $\min f(x)$  subject to  $g(x) \leq 0$  and  $h(x) = 0$ .

The gradient of the Lagrangian for (N.L.P.) is

$$\nabla L(x,u,w) := \nabla f(x) + \sum_{i=1}^m u_i \nabla g_i(x) + \sum_{j=1}^p w_j \nabla h_j(x)$$

As in Example 2, a Kuhn-Tucker triple for (N.L.P.) is a solution to the generalized equation (K.T.), defined by

$$(K.T.): \quad 0 \in F(z) + N_K(z)$$

where  $z := (x,u,w)$ ,  $K := R^n \times R_+^m \times R^p$ , and

$$F(z) := \begin{bmatrix} \nabla L(z) \\ -g(x) \\ -h(x) \end{bmatrix}.$$



Let  $G(x) := [\nabla g_1(x), \dots, \nabla g_m(x)]$  and  $H(x) := [\nabla h_1(x), \dots, \nabla h_p(x)]$

Then the derivative of  $F$  at  $\bar{z} := (\bar{x}, \bar{u}, \bar{w})$ ,  $F'(\bar{z})$ , is the matrix

$$F'(\bar{z}) = \begin{bmatrix} \nabla^2 L(\bar{z}) & G(\bar{x}) & H(\bar{x}) \\ -G^T(\bar{x}) & 0 & 0 \\ -H^T(\bar{x}) & 0 & 0 \end{bmatrix}, \text{ and the linearization of}$$

$F$  at  $\bar{z}$  is  $LF_{\bar{z}}(z) := F(\bar{z}) + F'(\bar{z})(z - \bar{z})$ . After some algebra, the linearization can be written as

$$LF_{\bar{z}}(z) = \begin{bmatrix} \nabla f(\bar{x}) \\ -g(\bar{x}) \\ -h(\bar{x}) \end{bmatrix} + \begin{bmatrix} \nabla^2 L(\bar{z}) & G(\bar{x}) & H(\bar{x}) \\ -G^T(\bar{x}) & 0 & 0 \\ -H^T(\bar{x}) & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x - \bar{x} \\ u \\ w \end{bmatrix}.$$

and the linearization of (K.T.) is  $0 \in LF_{\bar{z}}(z) + N_K(z)$ . Comparing this to the generalized equation for (Q.P.), we see that the linearization of the generalized equation (K.T.) has as its solution the Kuhn-Tucker triple for the quadratic program

$$\begin{aligned} \min \quad & \frac{1}{2}(x - \bar{x})^T \nabla^2 L(\bar{z})(x - \bar{x}) + (x - \bar{x})^T \nabla f(\bar{x}) \quad \text{subject to} \\ & Lg_{\bar{x}}(x) \leq 0 \quad \text{and} \quad Lh_{\bar{x}}(x) = 0. \end{aligned}$$

Thus, an iterative procedure which attempts to find the solution to a nonlinear programming problem by successively solving the linearized generalized equation will replace a general nonlinear programming problem with a sequence of quadratic programming problems.

4. Strong Regularity. An important element in the analysis of Newton's method for solving the equation  $0 = f(x)$ , where  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  has derivative  $f'$ , is the behavior of  $(f'(x))^{-1}$  for  $x$  in a neighborhood of a solution  $\bar{x}$ . The analogous element for generalized equations is the behavior of the set-valued map  $(Lf_x + N_K)^{-1}$  for  $x$  in a neighborhood of the solution  $\bar{x}$ . We begin with a definition.

Definition 6. Let  $T$  be a set-valued map.

Then

$$T^{-1}(y) := \{x | y \in T(x)\}$$

It follows that  $T^{-1}(0) = \{x | 0 \in T(x)\}$  is the solution set of the generalized equation  $0 \in T(x)$ . We are interested in  $T^{-1}$  when  $T = Lf_x + N_K$ . The notion of strong regularity, introduced in Robinson [9] describes the behavior of  $(Lf_x + N_K)^{-1}$  appropriate to the convergence of Newton's method.

Definition 7. (Robinson [9]). Let  $0 \in F(x) + N_K(x)$  have a solu-

tion  $z$ . Let  $T := LF_z + N_K$ .

Then  $z$  is a strongly regular solution of  $0 \in F(x) + N_K(x)$  if and only if there exists a neighborhood  $U$  of the origin, and a neighborhood  $W$  of  $z$ , such that the restriction of  $W \cap T^{-1}$  to  $U$  is a single-valued and Lipschitz continuous map.

The map  $W \cap T^{-1}$  is defined by  $(W \cap T^{-1})(y) := W \cap T^{-1}(y)$ . As noted by Robinson [9], if  $K = \mathbb{R}^n$ , then the generalized equation reduces to the equation  $0 = F(x)$ , and strong regularity is the condition that  $F'(z)$  has a continuous, linear inverse. Conditions under which the nonlinear complementarity problem of Example 1 and the nonlinear programming problem of Example 2 are strongly regular can be found in Robinson [9].

We conclude this section with a corollary to Theorem 2.4 of Robinson [9].

Corollary 1. Let  $C$  be a closed, convex, nonempty subset of  $\mathbb{R}^n$ , and let  $D$  be an open, convex, nonempty subset of  $\mathbb{R}^n$ . Let  $f: D \rightarrow \mathbb{R}^n$  have Fréchet derivative  $f'$ . Suppose  $x_0 \in D$  and the generalized equation  $0 \in Lf_{x_0}(x) + N_C(x)$  has a strongly regular solution  $x_1$  with associated Lipschitz constant  $d$ . Then there exist positive constants  $\rho$ ,  $r$ , and  $R$  such that the following holds:

For any  $x$  satisfying  $\|x - x_0\| \leq \rho$ ,

$(Lf_x + N_C)^{-1} \cap \bar{B}(x_1, r)$  restricted to  $\bar{B}(0, R)$  is single-

valued and Lipschitz continuous with modulus

$$d(1 - d \|f'(x) - f'(x_0)\|)^{-1},$$

where  $\bar{B}(x, s) := \{z \mid \|z - x\| \leq s\}$ .



5. Newton's Method for Generalized Equations. Newton's method for solving systems of equations is an iterative procedure in which each iteration involves the solution of a linear system of equations. There is an abundant literature on this technique. (See Ortega and Rheinboldt [4] and Rall [5] as representative samples of the literature). In this section, we will briefly recall Newton's method. We then prove two fundamental results on Newton's method for generalized equations, the domain of attraction theorem and the Newton-Kantorovich theorem.

Let  $D$  be an open, non-empty convex subset of  $\mathbb{R}^n$ , and let  $f:D \rightarrow \mathbb{R}^n$  be a continuously differentiable map. Newton's method for solving the equation  $0 = f(x)$  consists of replacing  $f$  by  $Lf_{\bar{x}}$  and solving  $0 = Lf_{\bar{x}}(x)$ . The point  $\bar{x}$  is some approximation to a zero of  $f$ . A solution  $\hat{x}$  to the linear system replaces  $\bar{x}$ , and the process is repeated (see Figure 1).

Newton's method for a generalized equation  $0 \in F(x) + N_K(x)$  consists of replacing  $F$  by  $LF_{\bar{x}}$  and solving the linearized equation  $0 \in LF_{\bar{x}}(x) + N_K(x)$ . The iteration procedure for generalized equation is formally identical to that in the case of a single-valued equation  $0 = f(x)$ .

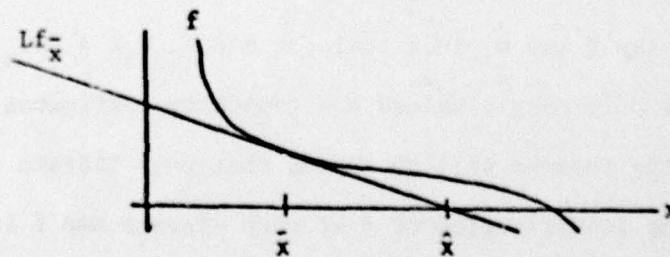


Figure 1. Newton Iterates

The domain of attraction theorem and the Newton-Kantorovich theorem are two fundamental results which describe the behavior of the Newton iteration process. The domain of attraction theorem assumes that a solution to an equation exists. Under conditions related to the behavior of the linearized equation in a neighborhood of the solution, a region is specified in which a vector used as an initial value for the iteration process can be chosen. Any such initial choice will lead to a well defined sequence which converges quadratically to the assumed solution.

The Newton-Kantorovich theorem is an existence result. Under conditions related to the behavior of the linearized equation in the neighborhood of an initial value, it states that the Newton iterates are well defined and converge to a solution of the original equation. The proof techniques for both theorems are adaptations to the generalized-equation case of classical techniques taken from Ortega and Rheinboldt [4].

Theorem 1 is the domain of attraction result. We assume  $f'$  is Lipschitz continuous on a region  $D$ . We also assume that the equation  $0 \in f(x) + N_C(x)$  has a strongly regular solution  $x^*$ . This implies the existence of a neighborhood of 0, say  $U$ , and two neighborhoods of  $x^*$ , say  $V$  and  $W$ , such that for  $x \in W$ ,  $T_x := (Lf_x + N_C)^{-1} \cap V$  restricted to  $U$  is single-valued and Lipschitz continuous. The hypotheses of the theorem will guarantee that each iterate remains in  $W \cap D$ , and the linearization of  $f$  at each iterate and  $f$  itself are close in value at  $x^*$ . This and the Lipschitz continuity of  $T_x$  will establish the theorem.



Theorem 2 is the Newton-Kantorovich theorem. The same assumptions are made on  $f$  and  $f'$ . The linearization of  $0 \in f(x) + N_C(x)$  at some initial point  $x_0$  is assumed to have a strongly regular solution  $x_1$ . The hypotheses of the theorem will guarantee properties of  $T_x$  similar to those in Theorem 1. The method of scalar majorants, described in detail in Ortega and Rheinboldt (4), is then invoked to establish the theorem. The major results of the method of majorants are summarized in Lemma A appearing in the appendix.

We will denote the closed ball about  $x$  with radius  $r$  by  $\bar{B}(x, r)$ .

Theorem 1. Let  $C$  be a nonempty, closed, convex subset of  $R^n$ , and let  $D$  be a nonempty, open, convex subset of  $R^n$ . Let  $f: D \rightarrow R^n$  have a Lipschitz continuous Frechet derivative  $f'$  with Lipschitz constant  $b$ . Suppose the generalized equation  $0 \in f(x) + N_C(x)$  has a strongly regular solution  $x^*$  with associated Lipschitz constant  $d$ . Let  $\rho$ ,  $r$ , and  $R$ , all positive, be the constants associated with strong regularity, as given in Corollary 1. Let  $x_0 \in D$ , and define the following notation.

$$m := \|x^* - x_0\|$$

$$S_x := Lf_x + N_C$$

$$a := d(1 - dbm)^{-1}$$

$$h := \frac{1}{2}abm$$

Suppose that the following relations hold:

$$(1) \quad dbm < \frac{2}{3}$$

$$(2) \quad \frac{1}{2}bm^2 < R$$

$$(3) \quad m < \rho$$

$$(4) \quad \bar{B}(x^*, m) \subset D \cap \bar{B}(x^*, \rho)$$



Then  $h < 1$ , the sequence given by

$$x_{n+1} := (S_{x_n})^{-1}(0) \cap \bar{B}(x^*, r)$$

is well defined and satisfies the following relations.

$$(5) \quad \|x_{n+1} - x^*\| \leq \frac{1}{2}ab \|x_n - x^*\|^2$$

$$(6) \quad \|x_{n+1} - x^*\| \leq 2 \cdot (ab)^{-1} h^{(2^{n+1})}$$

Proof. We first define the following notation.

$$T_x := \bar{B}(x^*, r) \cap S_x^{-1} \quad \text{and} \quad T_k := T_{x_k}$$

$$g(x) := d(1-d \|f'(x) - f'(x^*)\|)^{-1}$$

$$J(x) := Lf_x(x^*) - Lf_{x^*}(x^*)$$

We next establish the following relations.

$$(7) \quad \|T_x(u) - T_x(v)\| \leq a \|u - v\| \quad \text{for all } x \in \bar{B}(x^*, m) \text{ and } u, v \in \bar{B}(0, R),$$

$$(8) \quad x^* \in S_x^{-1}(J(x)) \quad \text{for all } x \in D,$$

$$(9) \quad x^* = T_x(J(x)) \quad \text{for } x \in \bar{B}(x^*, m) \text{ and } J(x) \in \bar{B}(0, R),$$

$$(10) \quad \|J(x)\| \leq \frac{b}{2} \|x - x^*\|^2.$$

Corollary 1 establishes that  $T_x$  is single-valued and Lipschitz continuous with modulus  $g(x)$  whenever  $x \in \bar{B}(x^*, \rho)$ . Also,

$$\|f'(x) - f'(x^*)\| \leq b \|x - x^*\| \leq bm \quad \text{for } x \in \bar{B}(x^*, m) \subset \bar{B}(x^*, \rho) \cap D,$$

thus  $g(x) \leq d(1-dbm)^{-1} =: a$ . Since  $g(x)$  is the Lipschitz constant for  $T_x$ , (7) is proven. Now let  $x \in D$ . Then we have

$$\begin{aligned} 0 \in f(x^*) + N_C(x^*) &= Lf_{x^*}(x^*) + N_C(x^*) \\ &= Lf_x(x^*) - J(x) + N_C(x^*) \\ &= -J(x) + S_x(x^*) + N_C(x^*) \end{aligned}$$

Thus,  $J(x) \in S_x(x^*)$ , which establishes (8). Equation (9) follows from (8) and  $T_x$  being single-valued on  $\bar{B}(x^*, m)$ . For  $x \in D$ , Lemma A.3 yields (10). Finally, we have

$$(11) \quad h < 1$$

since  $h = \frac{1}{2}abm = \frac{1}{2} \cdot dbm \cdot (1-dbm)^{-1} = \frac{1}{2} \frac{e}{1-e}$ , where  $e := dbm$ , and  $h < 1$  if and only if  $\frac{1}{2} \frac{e}{1-e} < 1$ , which occurs if and only if  $\frac{2}{3} > e := dbm$ , and this is hypothesis (1). The next part of the proof consists of using induction to prove that the following hold for all  $k \geq 1$ .

$$(12) \quad x_k := T_{k-1}(0) \text{ is well defined,}$$

$$(13) \quad \|x_k - x^*\| \leq \frac{1}{2}ab \|x_{k-1} - x^*\|^2 \leq hm$$

$$(14) \quad J(x_k) \in \bar{B}(0, R)$$

We first show that (12), (13) and (14) hold for  $k=1$ . By Corollary 1,  $T_0(0)$  is single-valued. Hence (12) holds. By (10) and (2),  $\|J(x_0)\| \leq \frac{b}{2} \|x_0 - x^*\|^2 = \frac{b}{2} m^2 < R$ , and thus, by (9),  $x^* = T_0(J(x_0))$ . We can now estimate the distance from  $x_1$  to  $x^*$ . We have

$$\|x_1 - x^*\| = \|T_0(0) - T_0(J(x_0))\| \leq a \|J(x_0)\| \quad \text{by (7)}$$

$$\leq \frac{1}{2}ba \|x_0 - x^*\|^2 = \frac{1}{2}abm^2 \quad \text{by (4) and (10)}$$

$$= hm < m \quad \text{by (11)}$$

Hence, by (3) and (4),  $x_1 \in \bar{B}(x^*, m) \subset D$ . We can also estimate the norm of  $J(x_1)$ . Using (10), we have

$$\begin{aligned}
\|J(x_1)\| &\leq \frac{b}{2} \|x_1 - x^*\|^2 \leq \frac{1}{2} b(hm)^2 = \frac{1}{2} b m^2 \cdot h^2 \\
&< \frac{1}{2} b m^2 && \text{by (11)} \\
&< R && \text{by (2).}
\end{aligned}$$

Thus, (12), (13), and (14) hold for  $k=1$ . Proceeding with the induction, we assume that (12), (13) and (14) hold for all  $k \leq n$  and show that they also hold for  $k=n+1$ . By (13), (3) and (11), it follows that (7) holds for  $x = x_n$ . By (14) and (9),  $x^* = T_n(J(x_n))$ . Thus,  $x_{n+1} = T_n(0)$  is well defined and

$$\begin{aligned}
\|x_{n+1} - x^*\| &= \|T_0(0) - T_n(J(x_n))\| \leq a \|J(x_n)\| && \text{by (7),} \\
&\leq \frac{1}{2} b a \|x_n - x^*\|^2 && \text{by (10), (13), and } \bar{B}(x^*, m) \subset D, \\
&\leq \frac{1}{2} b a (hm)^2 && \text{by (13)} \\
&= h^3 \cdot m < hm < m.
\end{aligned}$$

Also,

$$\begin{aligned}
\|J(x_{n+1})\| &\leq \frac{1}{2} b \|x_{n+1} - x^*\|^2 && \text{by (10) and } x_{n+1} \in \bar{B}(x^*, m) \subset D, \\
&\leq \frac{1}{2} b (hm)^2 \leq \frac{1}{2} b m^2 < R.
\end{aligned}$$

Thus, (12), (13), and (14) are established for  $k=n+1$ . This completes the induction. Relation (5) follows directly from (13), and (6) follows from (5) by induction. This completes the proof.

**Theorem 2.** Let  $C$ ,  $D$ ,  $f$ ,  $b$ , and  $S_x$  be as in Theorem 1. Let  $x_0 \in D$  and assume that the generalized equation  $0 \in S_{x_0}(x)$  is strongly regular at  $x_1$ . Let  $\rho$ ,  $r$ , and  $R$  be the positive constants associated



with strong regularity, and  $d$  the associated Lipschitz constant, as given in Corollary 1. Let  $m := \|x_1 - x_0\|$  and  $t^* := (db)^{-1}(1 - (1 - 2a)^{\frac{1}{2}})$ , where  $a := dbm$ . Suppose that the following relations hold:  $a \leq \frac{1}{2}$ ,  $\frac{1}{2}bm^2 \leq R$ , and  $\bar{B}(x_0, t^*) \subset D \cap \bar{B}(x_0, \rho)$ . Then the sequence  $\{x_n\}$  defined by

$$x_{n+1} := S_{x_n}^{-1}(0) \cap \bar{B}(x_1, r)$$

is well defined, converges to some  $x^* \in \bar{B}(x_0, t)$  which satisfies

$$0 \in f(x^*) + N_C(x^*), \text{ and } \|x^* - x_n\| \leq (2^n bd)^{-1} (2a)^{(2^n)} \text{ for } n=1, 2, \dots$$

Proof. We first define the following notation.

$$g(x) := d(1 - db\|x - x_0\|)^{-1},$$

$$G(t) := (\frac{1}{2}dbt^2 - m)(dbt - 1)^{-1},$$

$$Q(s, t) := (\frac{1}{2}dbs^2)(1 - dbt)^{-1},$$

$$t_0 := 0, \quad t_{k+1} := G(t_k) \text{ for } k \geq 0,$$

$$T_n := \bar{B}(x_1, r) \cap S_{x_n}^{-1},$$

$$Lf_n := Lf_{x_n},$$

$$J_n := f(x_n) - Lf_{n-1}(x_n).$$

Before beginning the induction part of the proof, we prove the following relation holds:

$$(15) \quad x_{n+1} = T_{n+1}(J_{n+1}) \text{ when } x_n, x_{n+1} \in \bar{B}(x_0, t^*),$$

$$x_n := T_n(0), \quad J_{n+1} \in \bar{B}(0, R) \text{ and } T_{n+1} \text{ restricted to}$$

$\bar{B}(0,R)$  is single-valued.

By definition of  $T_n$ ,  $x_{n+1} \in \bar{B}(x_1, r)$  and  $0 \in Lf_n(x_{n+1}) + N_C(x_{n+1})$ .

Thus

$$\begin{aligned} 0 &\in Lf_{n+1}(x_{n+1}) - (Lf_{n+1}(x_{n+1})) - Lf_n(x_{n+1}) + N_C(x_{n+1}) \\ &= Lf_{n+1}(x_{n+1}) - J_{n+1} + N_C(x_{n+1}) \end{aligned}$$

Rearranging results in  $J_{n+1} \in (Lf_{n+1} + N_C)(x_{n+1})$ , that is,

$x_{n+1} \in (Lf_{n+1} + N_C)^{-1}(J_{n+1})$ . Since  $J_{n+1} \in \bar{B}(0,R)$  and  $x_{n+1} \in \bar{B}(x_1, r)$ ,

we have  $x_{n+1} = T_{n+1}(J_{n+1})$ , which is (15). We now proceed by induction to prove the following for all  $n \geq 1$ .

(16)  $T_n$  restricted to  $\bar{B}(0,R)$  is single-valued and Lipschitz continuous with Lipschitz constant  $g(x_n)$ .

(17)  $\|x_{n+1} - x_n\| \leq Q(\|x_n - x_{n-1}\|, \|x_n - x_0\|)$ , where  $x_{n+1} := T_n(0)$ .

(18)  $x_{n+1} \in \bar{B}(x_0, t^*)$ .

(19)  $\|J_{n+1}\| \leq R$ .

We first show that (16) to (19) hold for  $n=1$ . By Lemma A.1 (4),

$t_1 \leq t^*$ . By definition,  $t_1 = G(t_0) = m = \|x_1 - x_0\|$ . Hence,

$$x_1 \in \bar{B}(x_0, t_1) \subset \bar{B}(x_0, t^*) \subset D \cap \bar{B}(x_0, \rho).$$

By Corollary 1,  $T_1$  restricted to  $\bar{B}(0,R)$  is a single-valued and Lipschitz continuous map with modulus  $d(1-d\|f'(x_1) - f'(x_0)\|)^{-1}$ .



By hypothesis,  $\|f'(x_1) - f'(x_0)\| \leq b \|x_1 - x_0\|$ , so that  $d(1 - d\|f'(x_1) - f'(x_0)\|)^{-1} \leq g(x_1)$ . Hence, (16) holds for  $n=1$ . Now define

$$x_2 = T_1(0). \text{ By Lemma A.3, } \|J_1\| := \|f(x_1) - Lf_0(x_1)\| \leq \frac{1}{2}b \|x_1 - x_0\|^2 = \frac{1}{2}bm^2 \leq R.$$

Thus,  $J_1 \in \bar{B}(0, R)$  and by (16),

$$\begin{aligned} \|T_1(J_1) - T_1(0)\| &\leq g(x_1) \cdot \|J_1\| \leq \frac{1}{2}db \|x_1 - x_0\|^2 (1 - db\|x_1 - x_0\|)^{-1} \\ &= Q(\|x_1 - x_0\|, \|x_1 - x_0\|). \end{aligned}$$

But  $x_2 = T_1(0)$  and by (15),  $x_1 = T_1(J_1)$ , hence we have

$$\|x_2 - x_1\| = \|T_1(0) - T_1(J_1)\|. \text{ This establishes (17) when } n=1. \text{ By}$$

Lemma A.2(6),  $x_2 \in \bar{B}(x_0, t^*)$ . Thus, (18) holds for  $n=1$ . Also,

$x_2 \in \bar{B}(x_0, t^*) \subset D$ , and thus

$$\begin{aligned} \|J_2\| &:= \|f(x_2) - Lf_1(x_2)\| \leq \frac{1}{2}b \|x_2 - x_1\|^2 \text{ by Lemma A.3} \\ &\leq \frac{1}{2}b(t_2 - t_1)^2 \text{ by Lemma A.2(9)} \\ &\leq \frac{1}{2}bm^2 2^{-1} \text{ by Lemma A.1(5)} \\ &\leq R \text{ by hypothesis.} \end{aligned}$$

Thus, (19) holds for  $n=1$ , and we have completed the induction when

$n=1$ . Suppose that (16) to (19) hold for all  $n \leq k$ . We proceed to show that they also hold for  $n=k+1$ . By (18),  $x_{k+1} \in \bar{B}(x_0, t^*) \subset \bar{B}(x_0, \rho) \cap D$ .



By Corollary 1,  $T_{k+1}$  restricted to  $\bar{B}(0, R)$  is single-valued and Lipschitz continuous with modulus  $d(1-d\|f'(x_{k+1}) - f'(x_0)\|)^{-1}$ .

By hypothesis,  $\|f'(x_{k+1}) - f'(x_0)\| \leq b\|x_{k+1} - x_0\|$ . Thus,

$g(x_{k+1}) \geq d(1-d\|f'(x_{k+1}) - f'(x_0)\|)^{-1}$ . Hence (16) holds for  $n=k+1$ .

Let  $x_{k+2} = T_{k+1}(0)$ . By (15) to (19),  $x_{k+1} = T_{k+1}(J_{k+1})$ . Thus,

$$\begin{aligned}\|x_{k+2} - x_{k+1}\| &= \|T_{k+1}(0) - T_{k+1}(J_{k+1})\| \leq g(x_{k+1})\|J_{k+1}\| \\ &\leq g(x_{k+1}) \cdot \frac{1}{2}b\|x_{k+1} - x_k\|^2 \\ &= Q(\|x_{k+1} - x_k\|, \|x_{k+1} - x_0\|).\end{aligned}$$

This establishes (17) for  $n = k+1$ . It follows from Lemma A.2(6) that

$x_{k+2} \in \bar{B}(x_0, t^*) \subset D$ , thus (18) holds for  $n=k+1$ . Now,

$$\begin{aligned}\|J_{k+2}\| &= \|f(x_{k+2}) - Lf_{k+1}(x_{k+2})\| \\ &\leq \frac{1}{2}b\|x_{k+2} - x_{k+1}\|^2 \quad \text{by Lemma A.3,} \\ &\leq \frac{1}{2}b(t_{k+2} - t_{k+1})^2 \quad \text{by Lemma A.2(9),} \\ &\leq \frac{1}{2}bm^2 \quad \text{by Lemma A.1(15),} \\ &\leq R\end{aligned}$$

Thus, (19) holds for  $n=k+1$ , which completes the induction. By Lemma

A.2, the sequence  $\{x_n\}$  converges to  $x^* \in \bar{B}(x_0, t^*)$  and

$\|x^* - x_n\| \leq (2^n bd)^{-1} (2a)(2^n)$ . It remains to show that  $x^*$  is a solution to the generalized equation  $0 \in f(x) + N_C(x)$ . Let graph  $N_C$

denote the graph of  $N_C$ , that is,  $\text{graph } N_C = \{(u,v) \mid v \in N_C(u)\}$ .

The assumptions on  $C$  imply that  $\text{graph } N_C$  is a closed subset of  $R^n \times R^n$ . By construction,

$$(x_{n+1}, -L f_n(x_{n+1})) \in \text{graph } N_C.$$

It follows from the continuity properties of  $f$  and  $f'$  on  $D$  that

$(x_{n+1}, -L f_n(x_{n+1}))$  converges to  $(x^*, -f(x^*))$ . Since  $\text{graph } N_C$  is closed,  $(x^*, -f(x^*))$  belongs to  $\text{graph } N_C$ , which implies

$0 \in f(x^*) + N_C(x^*)$ . This completes the proof of the theorem.



6. Computational Results. We present the results of applying Newton's method to two nonlinear complementarity problems.

Kojima [3] gives the following nonlinear complementarity problem:

$$f(x) \geq 0, \quad x \geq 0, \quad \langle x, f(x) \rangle = 0, \quad \text{where}$$

$$f(x) := \begin{bmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 3x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 1 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{bmatrix}$$

$$x := (x_1, x_2, x_3, x_4) \in \mathbb{R}^4.$$

The unique solution of the problem is given as

$$\hat{x}_1 = \sqrt{6}/2 = 1.2247449, \quad \hat{x}_2 = 0, \quad \hat{x}_3 = 0, \quad \hat{x}_4 = 0.5.$$

The generalized equation representing this nonlinear complementarity problem is

$$0 \in f(x) + N_{\mathbb{R}_+^4}(x).$$

The linearization of this generalized equation at the point  $\bar{x}$  is

$$0 \in f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + N_{\mathbb{R}_+^4}(x),$$

where

$$f'(\bar{x}) = \begin{bmatrix} 6\bar{x}_1 + 2\bar{x}_2 & 2\bar{x}_1 + 4\bar{x}_2 & 1 & 3 \\ 4\bar{x}_1 + 1 & 2\bar{x}_2 & 3 & 2 \\ 6\bar{x}_1 + \bar{x}_2 & \bar{x}_1 + 4\bar{x}_2 & 2 & 3 \\ 2\bar{x}_1 & 6\bar{x}_2 & 2 & 3 \end{bmatrix}.$$

If we let  $q := f(\bar{x}) - f'(\bar{x}) \cdot \bar{x}$  and  $M := f'(\bar{x})$ , then the linearization corresponds to the linear complementarity problem

$$x \geq 0, \quad Mx + q \geq 0, \quad \langle x, Mx + q \rangle = 0.$$



Lemke's complementary pivot algorithm (Cottle and Dantzig [1]) is applied to the above linear complementarity problem with the following results, as tabulated in Table 1. Three starting points, listed in column 1, all lead to a convergent sequence of linear problems. Column 2 lists the number of linear complementarity problems solved, with the final values of  $x$  and  $f$  listed in columns 3 and 4, respectively. Each linear complementarity problem produced a solution after two pivoting operations of the Lemke algorithm. Thus, for example, when the algorithm starts at  $x = (5,5,5,5)$ , there are six linear complementarity problems solved before the solution to the original nonlinear complementarity problem is found. Each linear complementarity problem is of size  $4 \times 4$ , and each is solved in two pivots of Lemke's algorithm. Thus, a total of six derivative evaluations and twelve pivot operations are required to solve this problem using Newton's method with initial condition  $x = (5,5,5,5)$ .

Table 1      Kojima Example

<u>Initial x</u>	<u>Iterations</u>	<u>Final x</u>	<u>Final f</u>
(1.,1.,1.,1.)	4	(1.2247449, 0, 0, .49999999)	( $.6 \cdot 10^{-7}$ , 3.224745, 5.000000, $-.3 \cdot 10^{-7}$ )
(3.,3.,3.,3.)	5	same	( $-.6 \cdot 10^{-7}$ , 3.224745, 4.9999999, $-.6 \cdot 10^{-7}$ )
(5.,5.,5.,5.)	6	same	( $-.1 \cdot 10^{-6}$ , 3.224775, 4.9999999, $-.8 \cdot 10^{-7}$ )

The second nonlinear complementarity problem is taken from Hansen and Koopmans [2]. Let  $v$  be a real-valued function on  $R_+^n$ , and consider the following maximization problem.

max  $v(x)$  subject to

$$Ax \leq z, \quad Bx \geq z, \quad Cx \leq w, \quad x \geq 0,$$

where  $w$  is a constant vector and  $z$  is fixed for the moment. The Kuhn-Tucker conditions for this problem are

$$d(x, q_1, q_2, r) := -\nabla v(x) + A^T q_1 - B^T q_2 + C^T r \geq 0, \quad x \geq 0, \quad \langle d(x, q_1, q_2, r), x \rangle = 0,$$

$$w - Cx \geq 0, \quad r \geq 0, \quad \langle w - Cx, r \rangle = 0,$$

$$z - Ax \geq 0, \quad q_1 \geq 0, \quad \langle z - Ax, q_1 \rangle = 0,$$

$$Bx - z \geq 0, \quad q_2 \geq 0, \quad \langle Bx - z, q_2 \rangle = 0,$$

where  $\nabla v$  denotes the gradient of  $v$ . Fix  $\alpha, 0 < \alpha < 1$ . The invariant stock problem of Hansen and Koopmans [2] consists of finding a vector  $z^*$  such that for some  $x^*, q_1^*, q_2^*$  and  $r^*$ ,

$$(i) \quad (x^*, q_1^*, q_2^*, r^*) \text{ satisfy the Kuhn-Tucker conditions when } z = z^*,$$

and

$$(ii) \quad q_2^* = \alpha q_1^*.$$

Replacing  $q_2$  by  $\alpha q_1$  in the Kuhn-Tucker conditions results in the following conditions which  $x^*, q_1^*$ , and  $r^*$  must satisfy. The subscript 1 on  $q_1$  has been dropped.

$$d^*(x, q, r) := -\nabla v(x) + (A^T - \alpha B^T)q + C^T r \geq 0, \quad x \geq 0, \quad \langle d^*(x, q, r), x \rangle = 0,$$

$$w - Cx \geq 0, \quad r \geq 0, \quad \langle w - Cx, r \rangle = 0,$$

$$(iii) \quad z - Ax \geq 0, \quad q \geq 0, \quad \langle z - Ax, q \rangle = 0,$$

$$(iv) \quad Bx - z \geq 0, \quad q \geq 0, \quad \langle Bx - z, q \rangle = 0.$$

However, conditions (iii) and (iv) can be replaced by  $(B-A)x \geq 0, \quad q \geq 0,$

$\langle (B-A)x, q \rangle = 0$ , since any  $x$  and  $q$  satisfying this condition will also satisfy (iii) and (iv) with  $z = Ax$ . The converse is immediate.

Thus, we have the following nonlinear complementarity problem.



$$d^*(x, q, r) \geq 0, \quad x \geq 0, \quad \langle d^*(x, q, r), x \rangle = 0,$$

$$w - Cx \geq 0, \quad r \geq 0, \quad \langle w - Cx, r \rangle = 0,$$

$$(B-A)x \geq 0, \quad q \geq 0, \quad \langle (B-A)x, q \rangle = 0.$$

The generalized equation representing this problem is

$$0 \in \begin{bmatrix} -\nabla v(x) \\ w \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & C^T & A^T - \alpha B^T \\ -C & 0 & 0 \\ B-A & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ r \\ q \end{bmatrix} + N_{R_+^s}(x, r, q)$$

where  $s$  is the sum of the dimensions of  $x, r$  and  $q$ . The linearization at  $(\bar{x}, \bar{r}, \bar{q})$  of this generalized equation is

$$0 \in \begin{bmatrix} -\nabla v(\bar{x}) + \nabla^2 v(\bar{x}) \cdot \bar{x} \\ w \\ 0 \end{bmatrix} + \begin{bmatrix} -\nabla^2 v(\bar{x}) & C^T & A^T - \alpha B^T \\ -C & 0 & 0 \\ B-A & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ r \\ q \end{bmatrix} + N_{R_+^s}(x, r, q),$$

which corresponds to the linear complementarity problem

$$Mx + t \geq 0, \quad x \geq 0, \quad \langle x, Mx + t \rangle = 0, \quad \text{where}$$

$$t := \begin{bmatrix} -\nabla v(\bar{x}) + \nabla^2 v(\bar{x}) \cdot \bar{x} \\ w \\ 0 \end{bmatrix}$$

and

$$M := \begin{bmatrix} -\nabla^2 v(\bar{x}) & C^T & A^T - \alpha B^T \\ -C & 0 & 0 \\ B-A & 0 & 0 \end{bmatrix}.$$

The results of applying Lemke's algorithm to this linear complementarity problem is summarized in Table 2. For the value of  $\alpha$  listed in column 1, each component of  $x$  is initialized to the value given in column 2. The number of iterations performed to reach the solution published in Hansen and Koopmans [2, Table 2] is listed in column 3, and the average number



of pivots needed in Lemke's algorithm to compute each iterate is given in column 4. It should be noted that potential difficulties exist in applying Newton's method to this problem, since the derivative  $\nabla v$  does not exist for certain values of  $\bar{x}$ . However, this difficulty did not arise for the initial conditions chosen as test cases.

Table 2      Hansen-Koopmans Example

<u><math>\alpha</math></u>	<u>Initial <math>x</math></u>	<u>Iterations</u>	<u>Av. Pivots per It.</u>
.7	.1	8	10
.7	.2	14	10
.7	.5	14	10
.8	.3	11	10
.9	.3	8	14

# APPENDIX A

## RESULTS FROM ORTEGA AND RHEINBOLDT [4]

The following two results can be found in Chapter 12 of Ortega and Rheinboldt ( 4 ). The third result is standard.

Lemma A.1. Let  $d > 0$ ,  $b > 0$  and  $m > 0$ . Let  $a := bdm$  and suppose  $a \leq \frac{1}{2}$ . Let  $t^* := (bd)^{-1}(1 - (1-2a)^{\frac{1}{2}})$  and  $t^{**} := (bd)^{-1}(1 + (1-2a)^{\frac{1}{2}})$ . Define  $t_0 := 0$  and  $t_{k+1} := G(t_k)$ ,  $k \geq 0$ , where  $G(t) := (\frac{1}{2}dbt^2 - m)(dbt - 1)^{-1}$ . Let  $Q(s, t) := (\frac{1}{2}dbs^2)(1 - dbt)^{-1}$ . Then the following hold.

- (1)  $t_{k+1} = t_k + Q(t_k - t_{k-1}, t_k)$  for all  $k \geq 1$ .
- (2)  $G$  is monotone non-decreasing on  $[0, t^*]$ .
- (3)  $t^*$  is the smallest fixed point of  $G$  on  $[0, t^{**}]$ .
- (4)  $\lim_{k \rightarrow \infty} t_k = t^*$ , and the sequence  $\{t_k\}$  is monotone increasing.
- (5)  $t_{k+1} - t_k \leq m \cdot 2^{-k}$  for all  $k \geq 0$ .
- (6)  $t_{k+1} \leq 2[1 - (\frac{1}{2})^{k+1}]m$  for all  $k \geq 0$ .
- (7)  $|t^* - t_k| \leq (2a)^{(2^k)} \cdot (bd \cdot 2^k)^{-1}$  for all  $k \geq 0$ .

Lemma A.2. Let  $D$  be a subset of  $R^m$  and let  $\{x_n\}$ ,  $n \geq 0$ , be a sequence in  $R^m$ . Assume that the following conditions hold, where  $Q$ ,  $\{t_k\}$ ,  $k \geq 0$ ,  $t^*$ ,  $a$ ,  $d$ , and  $b$  are as in Lemma A.1.



- (1)  $\|x_{n+2} - x_{n+1}\| \leq Q(\|x_{n+1} - x_n\|, \|x_{n+1} - x_0\|)$  whenever  $x_n, x_{n+1} \in D$ .
- (2)  $x_0 \in D$ .
- (3)  $t_1 \geq \|x_1 - x_0\|$ .
- (4)  $x \in D$  whenever  $\|x - x_0\| \leq t^*$ .
- (5) The hypotheses of Lemma A.1. hold.

Then the following conclusions are valid:

- (6)  $\|x_n - x_0\| \leq t^*$  for all  $n \geq 0$ .
- (7)  $\{x_n\}$  converges to some  $x^*$  with  $\|x^* - x_0\| \leq t^*$ .
- (8)  $\|x^* - x_{k+1}\| \leq t^* - t_{k+1} \leq (2a)(2^{k+1})(db \cdot 2^{k+1})^{-1}$  for all  $k \geq 0$ .
- (9)  $\|x_{n+1} - x_n\| \leq t_{n+1} - t_n$  for all  $n \geq 0$ .

The following standard result can be found in Ortega and Rheinboldt [4].

Lemma A.3. Let  $D$  be an open, convex subset of  $R^n$ . Let  $f: D \rightarrow R^n$  be continuously differentiable. Suppose that for some  $k > 0$ ,  $\|f'(u) - f'(v)\| \leq k\|u - v\|$  whenever  $u, v \in D$ . Then  $\|f(u) - Lf_v(u)\| = \frac{1}{2}k\|u - v\|^2$  whenever  $u, v \in D$ .



#### ACKNOWLEDGEMENTS

Parts of the research reported in this paper are based on the author's doctoral thesis submitted to the Department of Industrial Engineering at the University of Wisconsin-Madison, written under the direction of Professor Stephen M. Robinson.

I wish to express my sincere gratitude to Professor Stephen M. Robinson for his advice and support. His continued encouragement, guidance and understanding throughout my graduate studies have been invaluable to me.

NHJ:db

# REFERENCES

- [1] Cottle, R. W. and Dantzig, G. B. (1967). Complementary Pivot Theory of Mathematical Programming. Technical Report 67-2, Operations Research House, Stanford University, California.
- [2] Hansen, T. and Koopmans, T. C. (1972). On the Definition and Computation of a Capital Stock Invariant under Optimization. J. Econ. Theory 5 , 487-523.
- [3] Kojima, M. (1974). Computational Methods for Solving the Nonlinear Complementarity Problem. Keio Engineering Reports, 27 , 1-41.
- [4] Ortega, J. M. and Rheinboldt, W. C. (1970). Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York.
- [5] Rall, L. B. (1969). Computational Solution of Nonlinear Operator Equations. Wiley, New York.
- [6] Robinson, S. M. (1976). An Implicit-Function Theorem for Generalized Variational Inequalities. Technical Summary Report No. 1672, Mathematics Research Center, University of Wisconsin, Madison.
- [7] Robinson, S. M. (1976). Stability Theory for Systems of Inequalities, Part II: Differentiable Nonlinear Systems. SIAM J. Num. Anal. 13, 497-513.
- [8] Robinson, S. M. (1976). Regularity and Stability for Convex Multi-valued Functions. Math. of Operations Research 1, 130-143.
- [9] Robinson, S. M. (1978). Strongly Regular Generalized Equations. Technical Summary Report No. 1877, Mathematics Research Center, University of Wisconsin, Madison.
- [10] Robinson, S. M. (1979). Generalized Equations and Their Solutions, Part I: Basic Theory. Math. Programming Studies 10, 128-141.
- [11] Wilson, R. B. (1963). A Simplicial Algorithm for Concave Programming. Dissertation, Graduate School of Business Administration, Harvard University, Boston, Massachusetts.



REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM	
1. REPORT NUMBER 1965	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) NEWTON'S METHOD FOR GENERALIZED EQUATIONS.		5. TYPE OF REPORT & METHOD COVERED Technical Summary Report, no specific reporting period	
7. AUTHOR(s) 10 Norman H./Josephy		8. CONTRACT OR GRANT NUMBER(s) DCR 74-20584 DAAG29-75-C-0024 KASE-MCS 74-20584 A02	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 5 - Mathematical Programming and Operations Research	
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below		12. REPORT DATE June 1979	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 14 MRC-TSR-1965		13. NUMBER OF PAGES 32	
		15. SECURITY CLASS. (of this report) UNCLASSIFIED	
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.			
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)			
18. SUPPLEMENTARY NOTES U. S. Army Research Office P.O. Box 12211 Research Triangle Park North Carolina 27709 National Science Foundation Washington, D. C. 20550			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Variational Inequalities, Complementarity, Newton's method			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Newton's method is a well known and often applied technique for computing a zero of a nonlinear function. By using the theory of generalized equations, a Newton method is developed to solve problems arising in both mathematical programming and mathematical economics. We prove two results concerning the convergence and convergence rate of Newton's method for generalized equations. Examples are given to emphasize the application of this method to generalized equations representing the nonlinear programming problem and the nonlinear complementarity problem. (continued)			



Abstract (continued)

We present computational results of Newton's method applied to a nonlinear complementarity problem of Kojima [3] and an invariant capital stock problem of Hansen and Koopmans [2].